A Junction Tree based Reinforcement Learning Algorithm for Coordinated Multi Agent Systems to Solve Network level Signal Control Problems

Paper No. 14-3649

Feng Zhu (zhu214@purdue.edu), H. M. **Abdul** Aziz (haziz@purdue.edu), Xinwu Qian (xinwu.qian.1@purdue.edu); Purdue University
Satish V. Ukkusuri, Associate Professor, Purdue University (sukkusur@purdue.edu)

Introduction

Optimized traffic control systems :
-Travel time reduction
-Saving in fuel consumption
-Cut down emissions
Traditional control :
SCOOT (Hunt et al., 1982), SCATS (Lowrie, 1982), PRODYN (Farges et al., 1983), OPAC (Gartner, 1983), RHODES (Mirchandani and Head, 2001), UTOPIA (Mauro and Taranto, 1989)
Limitation:
- Isolated learning, no network wide coordination
- No consideration of dynamic feedback

Research objectives

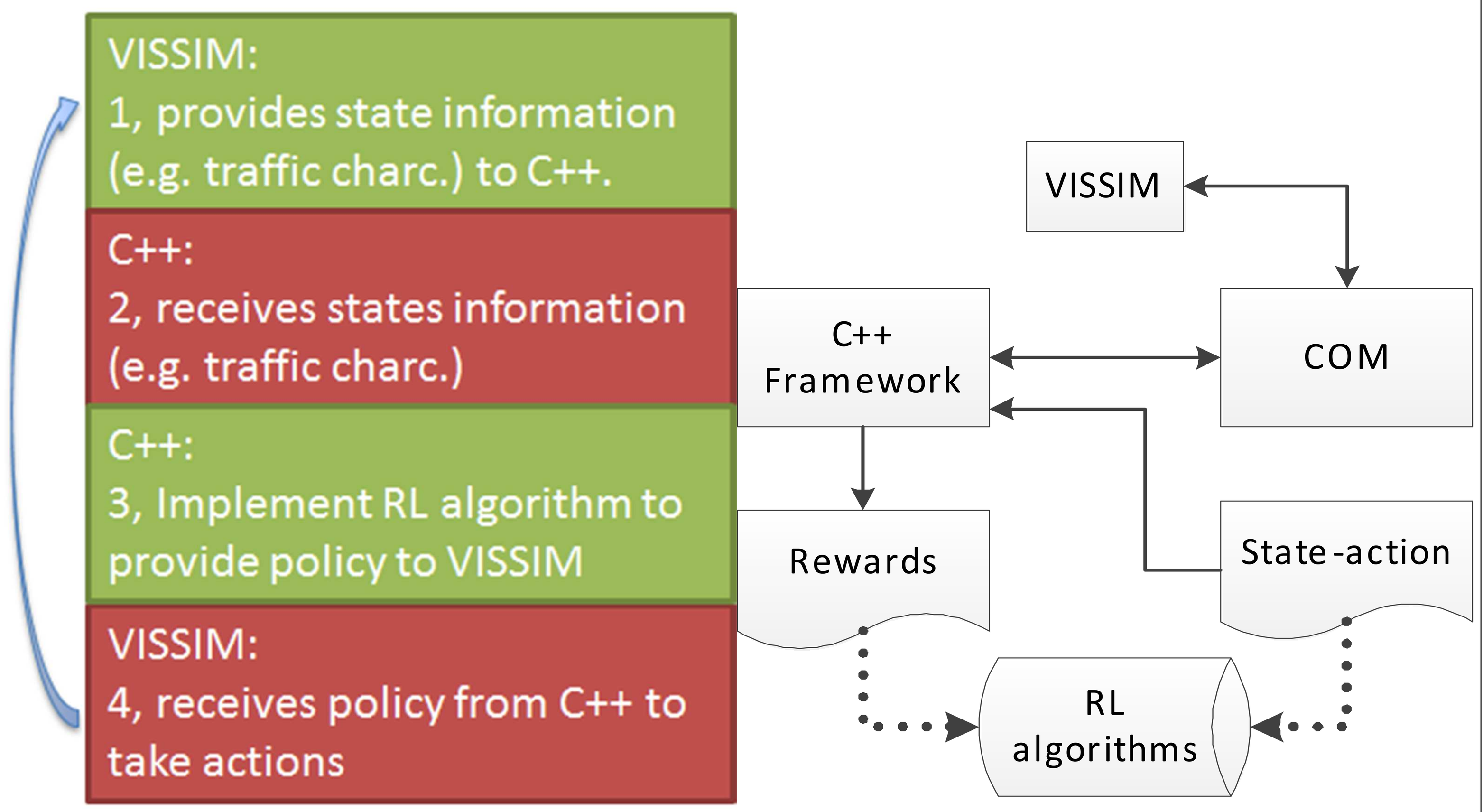
1) To propose multi-agent RL based signal control algorithm where agents coordinate their decisions
2) Compare the coordinated learning and isolated learning algorithms
3) To demonstrate the coordinated control algorithm as a potential application in the CV environment
4) To assess the environmental impacts of the proposed controller using a dynamic emissions simulator (MOVES2010)

Problem description

- o ***Major assumption*** :
- The intersections coordinate with each other to benefit the entire system.
- o ***Signal control*** :
- To find the optimal policy (mapping between the phase activations and traffic states)
- o ***Develop a coordinated Reinforcement Learning algorithm*** :
- Traffic lights in the network act as agents and take actions (activating the phase) according to RL algorithm.

Methodology

o ***Modeling Framework***



Junction Tree Algorithm

Formulation :

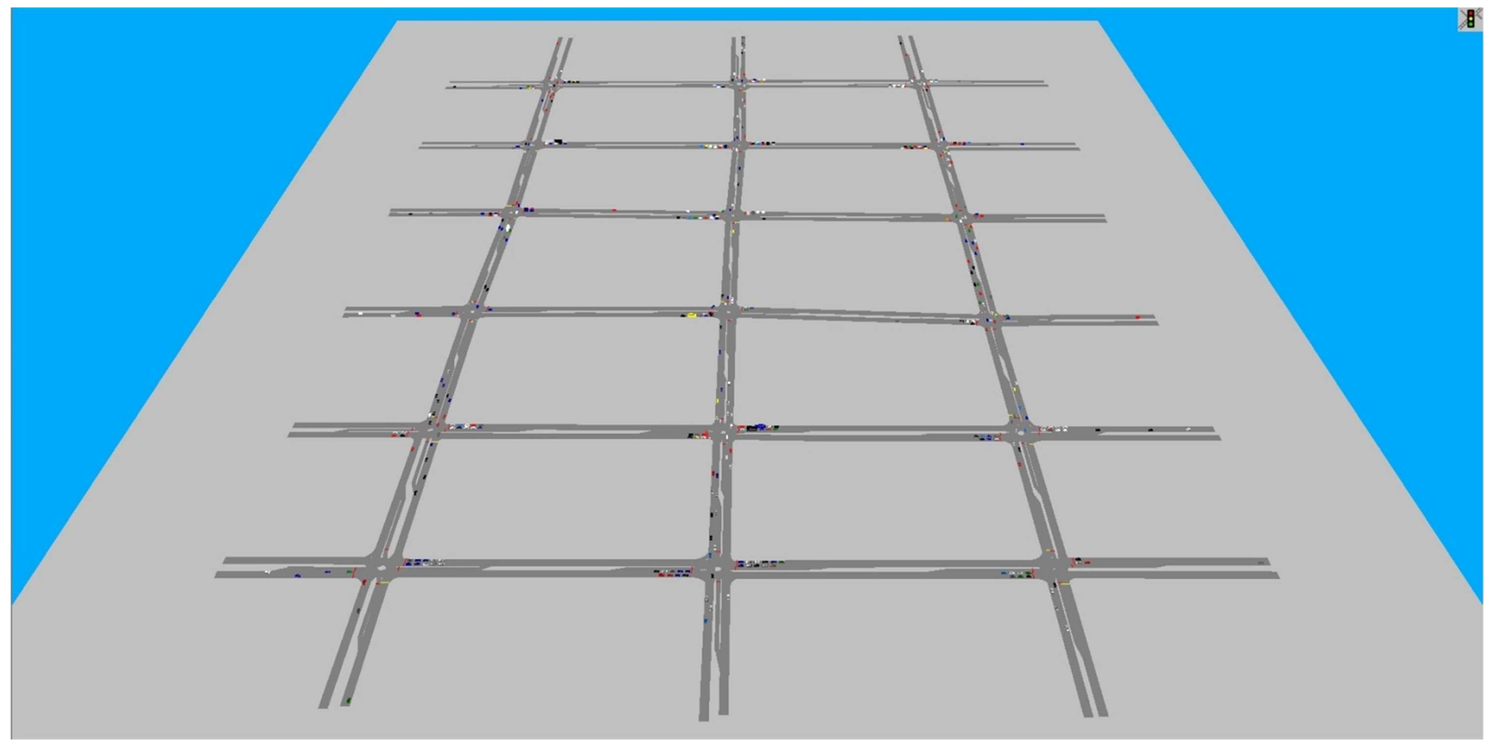
$$\max Q = max \sum_{i=1}^{N} Q_i$$

ϕ_i: potential function
Q_i: local Q-values

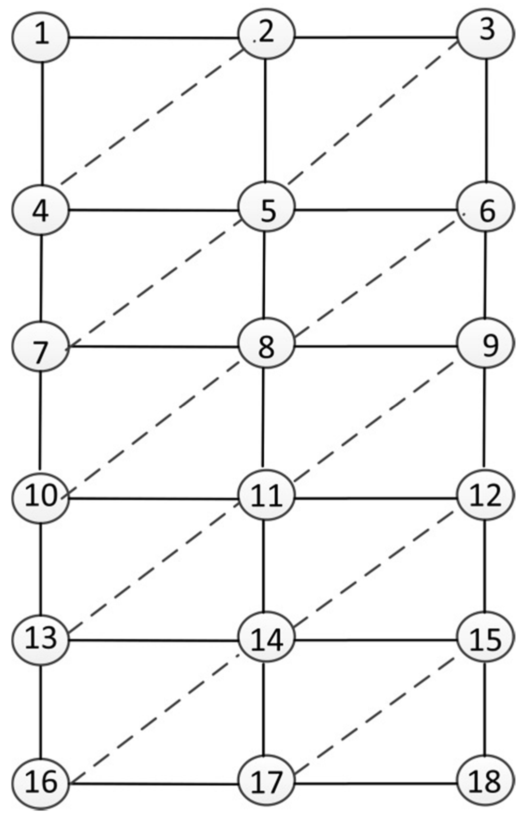
Five steps :

a. ***Moralization:*** converts a directed graph into an undirected graph
b. ***Introduction of potential:*** a new term to describe the characterization of the cluster
c. ***Triangulation:*** ensure that every loop of length 4 or more has a chord
d. ***Construction of junction tree:*** find out the maximal spanning tree
e. ***Propagation of messages:*** forward and backward message passing

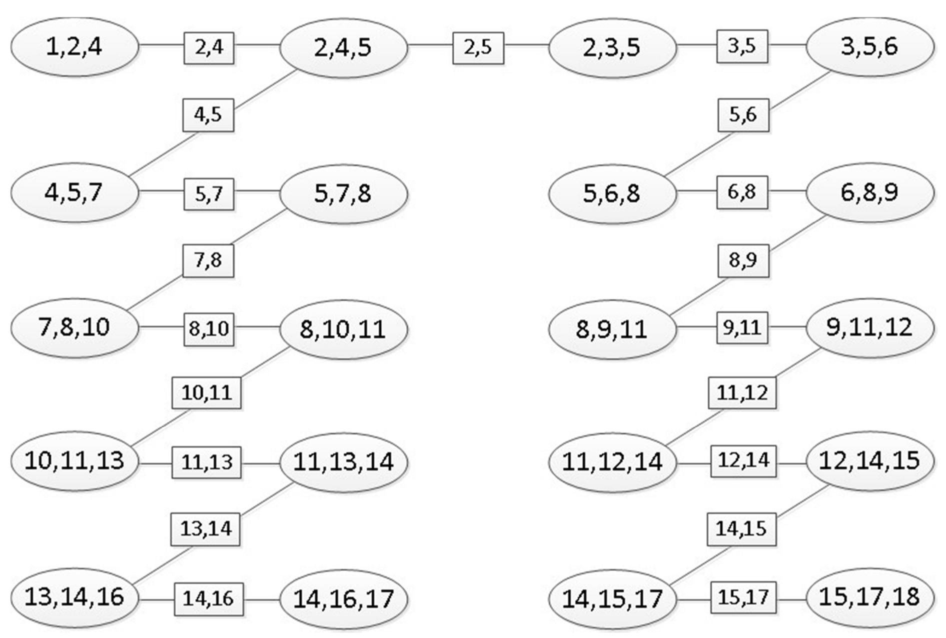
Test case study



(a) Original network



(b) Triangulated network



(c) Junction tree network

Experiment design

Demand scenarios :
- Low (500vph~600vph)
- Medium (600vph~900vph)
- High (900vph~1200vph)

Comparison of algorithms
- JTA: junction tree algorithm
- Max-plus: coordinated RL algorithm (Kuyer et al., 2008)
- Q-learning: independent RL algorithm (Abdulhai et al., 2003)
- LQF: longest queue first algorithm (Wunderlich et al., 2008)
- Fixed-control: fixed time signal control

Result analysis

Comparison between JTA and Q-learning

Congestion level	Algorithm	Delay (in seconds)	Stopped delay (in seconds)	Stops
Low	Q-learning	6.81	3.59	0.48
	Max-plus	6.55	3.31	0.48
	JTA	6.24	3.03	0.47
Medium	Q-learning	9.60	5.60	0.58
	Max-plus	6.36	2.89	0.48
	JTA	6.33	2.88	0.48
High	Q-learning	14.98	11.42	0.56
	Max-plus	11.83	5.90	0.79
	JTA	11.22	5.46	0.78

Comparison between JTA and LQF

Congestion level	Algorithm	Delay (in seconds)	Stopped delay (in seconds)	Stops
	LQF	7.07	3.99	0.44
Low	Fixed timing	15.79	11.93	0.58
	JTA	6.24	3.03	0.47
	LQF	10.34	6.45	0.54
Medium	Fixed timing	17.15	12.60	0.68
	JTA	6.33	2.88	0.48
	LQF	14.93	9.55	0.71
High	Fixed timing	25.67	17.56	1.32
	JTA	11.22	5.46	0.78

Total emissions for two intersections in the network

Pollutant	JTA	Q-Learning	LQF	Fixed control
CO (g/hour)	26363	26285	28570	37263
CO_2 (kg/hour)	2293	2447	2644	3204
NO_x (g/hour)	1799	1689	1857	2589
VOC(g/hour)	869	1029	1073	1186
PM_{10}	133	151	144	185

Conclusions

- Development of a coordinated RL based algorithm for signal control
- Junction tree algorithm to obtain best joint actions
- Explore the environmental benefits of the algorithm
- Test results show significant advantages of coordinated learning over independent learning of agents.